Using Mathematical Modelling to Understand the Role of Linker Histone Dynamics in DNA Packaging

Gustavo Carrero

Athabasca University

International Interdisciplinary Science Conference on Protein Folding and Diseases, New Delhi, India, December 8, 2012

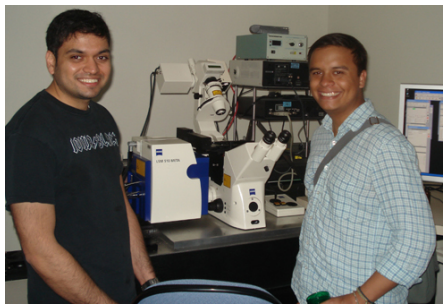
Athabasca University
CANADA'S OPEN UNIVERSITY

UNIVERSIDAD SIMÓN BOLÍVAR

UNIVERSITY OF ALBERTA

Acknowledgements

- Nikhil Raghuram and Carlos Contreras *(University of Alberta and Universidad Simón Bolívar)*



UNIVERSITY OF ALBERTA

UNIVERSIDAD SIMÓN BOLÍVAR

- Minalla Villasana *(Universidad Simón Bolívar)*

UNIVERSIDAD SIMÓN BOLÍVAR

- Michael Hendzel *(University of Alberta)*

UNIVERSITY OF ALBERTA

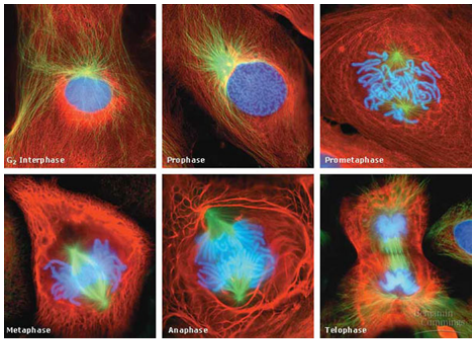
- Athabasca University Research Incentive Grant (AU-RIG)

Athabasca University

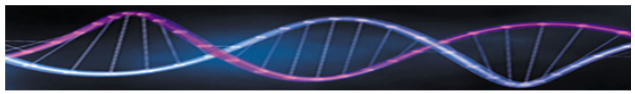
Outline

Histone H1 or Linker Histones

- Length of all DNA in an adult human cell $\sim$ 2 mts



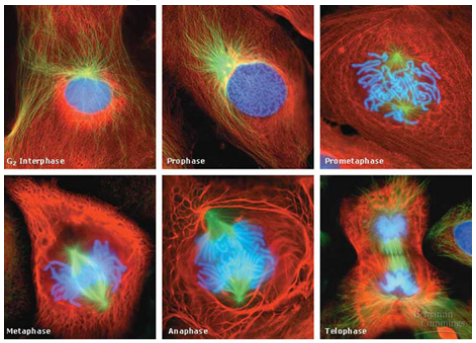
x 1000 times



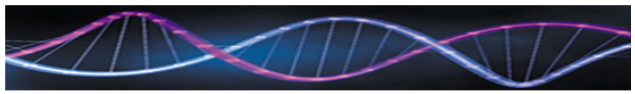
2 kms of DNA

Histone H1 or Linker Histones

- Length of all DNA in an adult human cell $\sim$ 2 mts

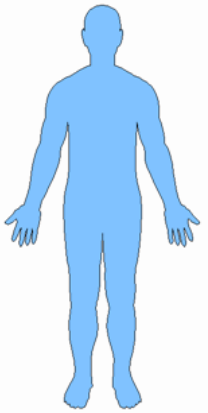


x 1000 times

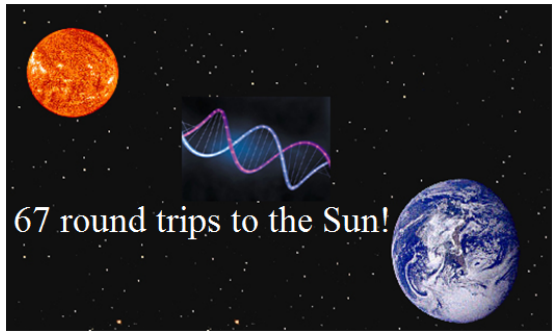


2 kms of DNA

- There are $\sim$ 10 trillion (10×10^{13}) cells in the body



~20 trillion mts. of DNA
in the human body



Histone H1 or Linker Histones

How do 133 AU of DNA fit in our body?

How is the DNA packed and organized in the cells?

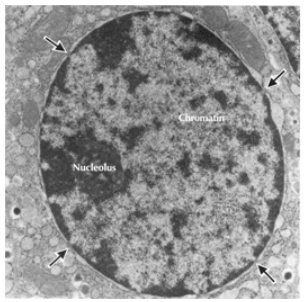
Histone H1 or Linker Histones

How do 133 AU of DNA fit in our body?

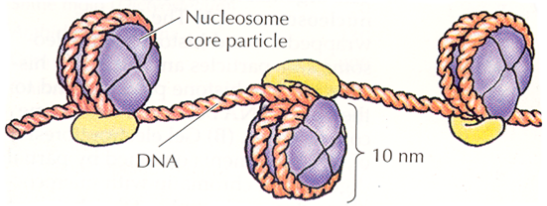
How is the DNA packed and organized in the cells?

With the help of histones!

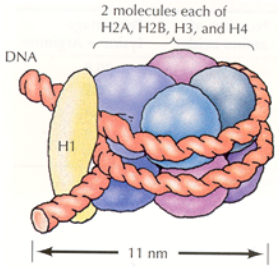
DNA is wrapped around core histones and locked by linker histones.



Chromatin distribution in the nucleus

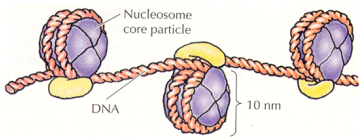


Chromatin: DNA and associated proteins

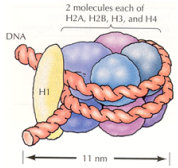


Nucleosome: Core Histones and Linker Histone

Histone H1 or Linker Histones



Chromatin: DNA and associated proteins



Nucleosome:
Core Histones and
Linker Histone

1884: Discovery of histones (Albrecht Kossel)

↓

$>$ 100 years: Histone is just *inert packing material*

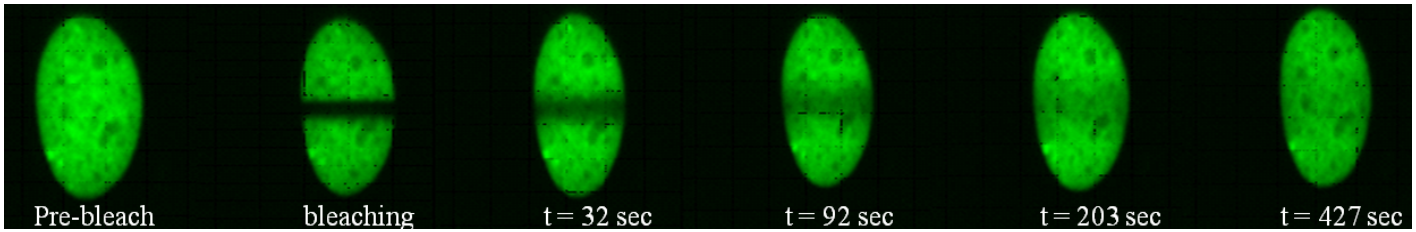
↓

1990's: Histone is a *regulator of gene expression!*

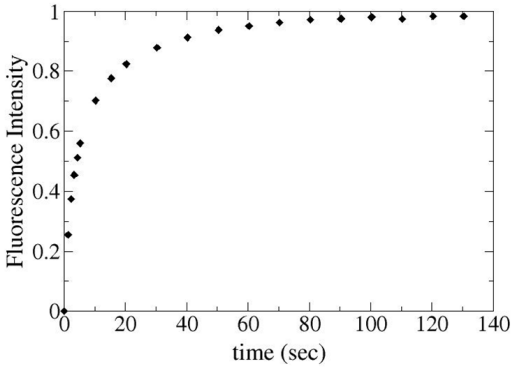
↓

2000's: Histone is a *dynamic regulator of gene expression!*

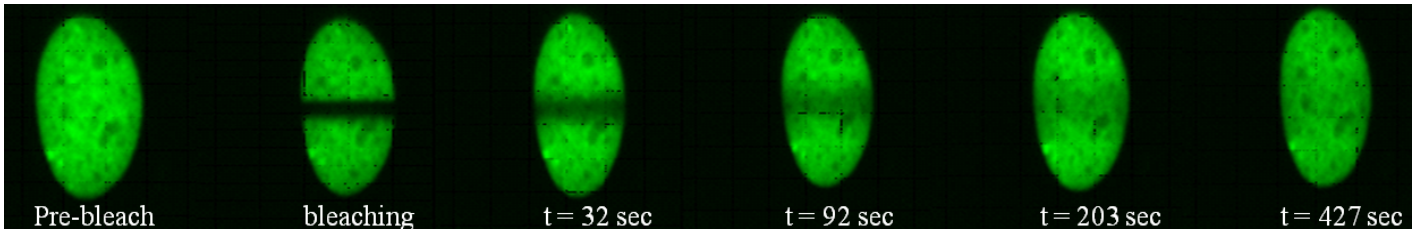
Mobility of Nuclear Proteins and FRAP Experiments
(Phair & Misteli, 2000; Kruhlak et al., 2000, Hendzel et al., 2000)



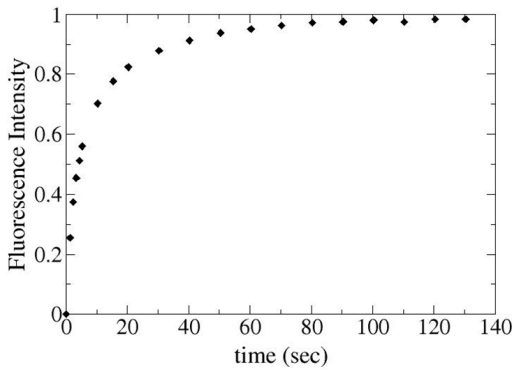
GFP (Green Fluorescence Protein)



Mobility of Nuclear Proteins and FRAP Experiments
(Phair & Misteli, 2000; Kruhlak et al., 2000, Hendzel et al., 2000)



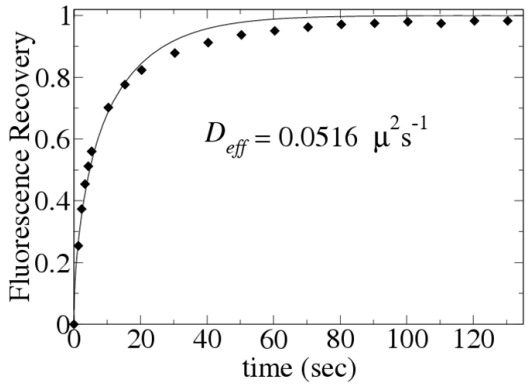
GFP (Green Fluorescence Protein)



$$\frac{\partial}{\partial t}\, u(x,t) = D_{eff}\, \frac{\partial^2}{\partial x^2}\, u(x,t)$$

$$R(t; D_{eff}) = \int_{\Lambda} u(x,t)\, dx$$

Λ =photobleached region

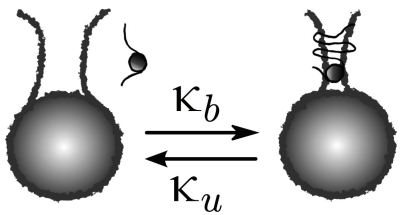


Mathematical Modelling and FRAP within the Cell Nucleus
(Carrero et al., 2003)

Outline

A dynamical model for histone H1

(Carrero et al., 2004)



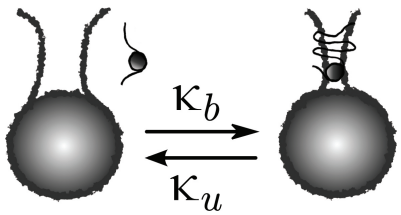
$$\frac{\partial}{\partial t}\, u(x,t) = D\, \frac{\partial^2}{\partial x^2}\, u(x,t) - k_b u(x,t) + k_u v(x,t)\ ,$$

$$\frac{\partial}{\partial t}\, v(x,t) = k_b u(x,t) - k_u v(x,t)$$

- Two populations of histone H1:
  - Unbound: A freely diffusing population u.
  - Bound: A population v bound to the chromatin structure.
- Parameters to be estimated: D, k_b, k_u

A binding model for histone H1

(3 parameters: D, k_b, k_u)

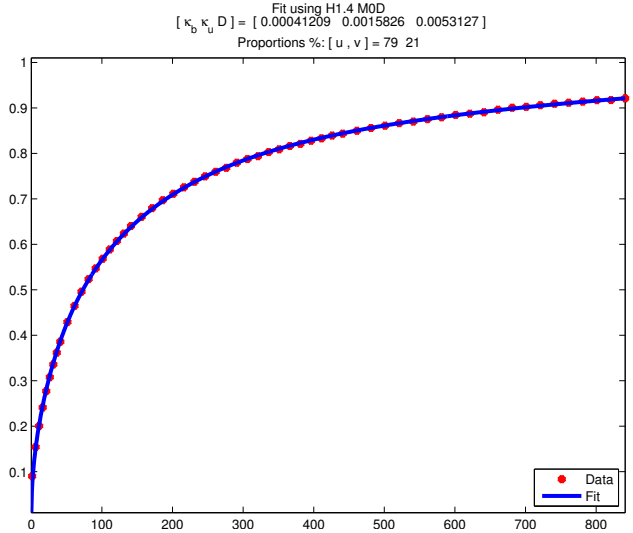


$$\frac{\partial}{\partial t}\, u(x,t) = D\, \frac{\partial^2}{\partial x^2}\, u(x,t) - k_b u(x,t) + k_u v(x,t)\ ,$$

$$\frac{\partial}{\partial t}\, v(x,t) = k_b u(x,t) - k_u v(x,t)$$

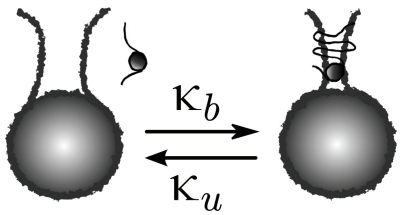
$$R(t; D, k_b, k_u) = \int_\Lambda [u(x,t) + v(x,t)]\, dx$$

- Proportion of bound population:
  $P_b = \dfrac{k_b}{k_b + k_u} \approx 21\%$ (too low!)
- Proportion of unbound population:
  $P_u = \dfrac{k_u}{k_b + k_u} \approx 79\%$
- Diffusion Coefficient:
  $D \approx 5 \times 10^{-3} \mu\mathrm{m}^2/\mathrm{sec}$ (too small!)



A binding model for histone H1

(2 parameters: k_b, k_u; D is fixed)



$$\frac{\partial}{\partial t}\, u(x,t) = D\, \frac{\partial^2}{\partial x^2}\, u(x,t) - k_b u(x,t) + k_u v(x,t)\ ,$$

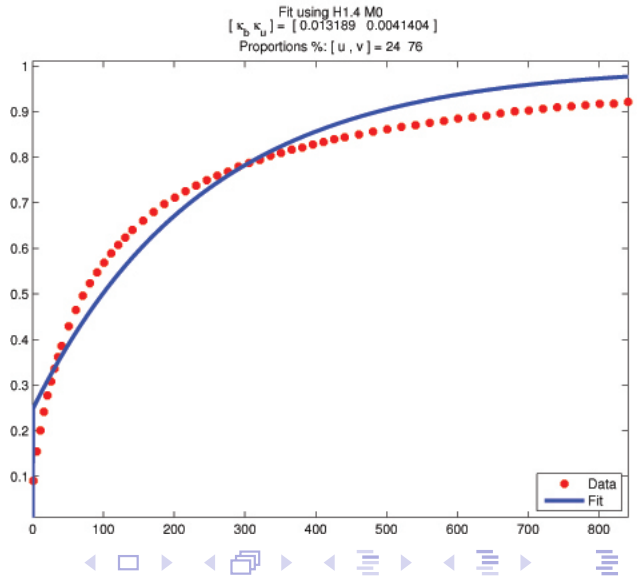
$$\frac{\partial}{\partial t}\, v(x,t) = k_b u(x,t) - k_u v(x,t)$$

$$R(t; k_b, k_u) = \int_\Lambda [u(x,t) + v(x,t)]\, dx$$

Fixed Diffusion Coefficient: $D = 40 \mu \mathrm{m}^2/\mathrm{sec}$

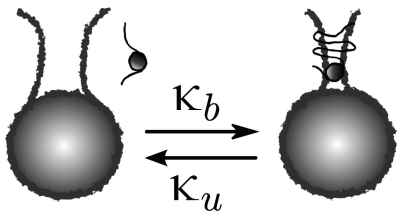
- Proportion of bound population:
  $$P_b = \frac{k_b}{k_b + k_u} \approx 76\%$$
- Proportion of unbound population:
  $$P_u = \frac{k_u}{k_b + k_u} \approx 24\%$$

▶ Fitting looses accuracy!



A binding model for histone H1

(3 parameters: D, k_b, k_u)

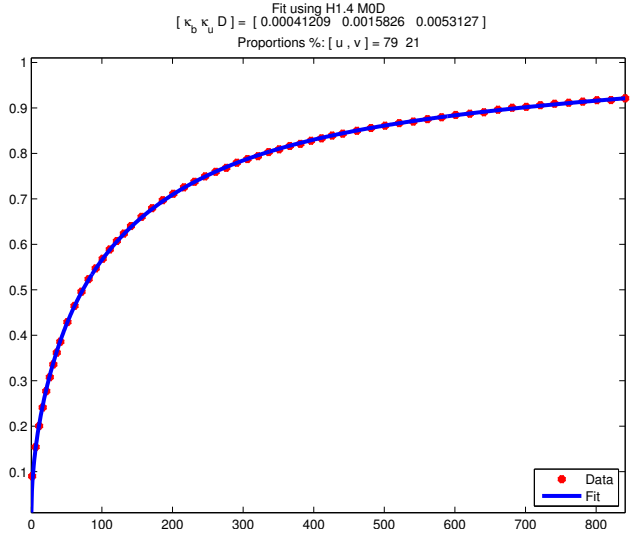


$$\frac{\partial}{\partial t}\, u(x,t) = D\, \frac{\partial^2}{\partial x^2}\, u(x,t) - k_b u(x,t) + k_u v(x,t)\ ,$$

$$\frac{\partial}{\partial t}\, v(x,t) = k_b u(x,t) - k_u v(x,t)$$

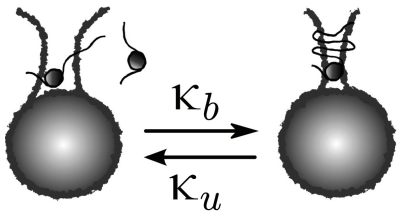
$$R(t; D, k_b, k_u) = \int_\Lambda [u(x,t) + v(x,t)]\, dx$$

- Proportion of bound population:
  $P_b = \dfrac{k_b}{k_b + k_u} \approx 21\%$ (too low!)
- Proportion of unbound population:
  $P_u = \dfrac{k_u}{k_b + k_u} \approx 79\%$
- Diffusion Coefficient:
  $D \approx 5 \times 10^{-3} \mu\mathrm{m}^2/\mathrm{sec}$ (too small!)



A binding model for histone H1

(3 parameters: D_{eff}, k_b, k_u)

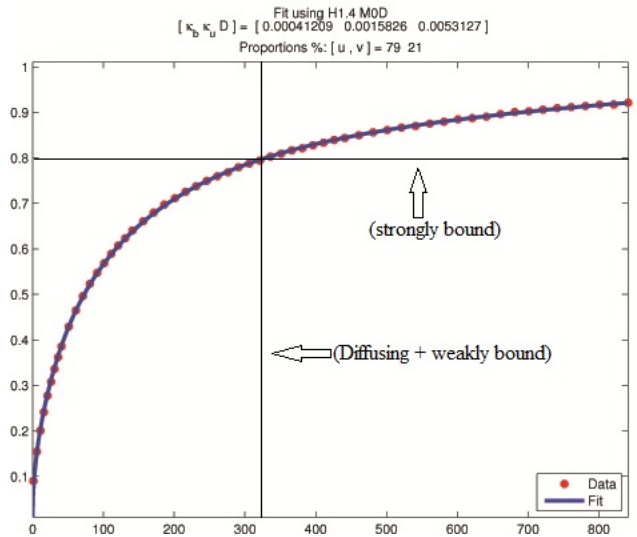


$$\frac{\partial}{\partial t}\, u(x,t) = D_{eff}\, \frac{\partial^2}{\partial x^2}\, u(x,t) - k_b u(x,t) + k_u v(x,t)\ ,$$

$$\frac{\partial}{\partial t}\, v(x,t) = k_b u(x,t) - k_u v(x,t)$$

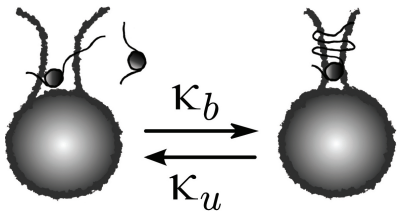
$$R(t; D_{eff}, k_b, k_u) = \int_\Lambda [u(x,t) + v(x,t)]\, dx$$

- Proportion of strongly bound population:
  $P_b = \dfrac{k_b}{k_b + k_u} \approx 21\%$
- Proportion of weakly bound and unbound population:
  $P_u = \dfrac{k_u}{k_b + k_u} \approx 79\%$
- Effective Diffusion Coefficient:
  $D_{eff} \approx 5 \times 10^{-3} \mu\mathrm{m}^2/\mathrm{sec}$



A binding model for histone H1

(3 parameters: D_{eff}, k_b, k_u)



$$\frac{\partial}{\partial t}\, u(x,t) = D_{eff}\, \frac{\partial^2}{\partial x^2}\, u(x,t) - k_b u(x,t) + k_u v(x,t)\ ,$$

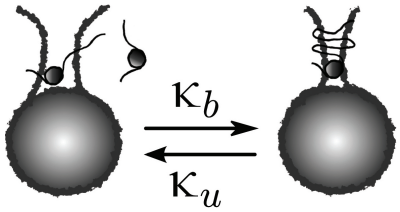
$$\frac{\partial}{\partial t}\, v(x,t) = k_b u(x,t) - k_u v(x,t)$$

$$R(t; D_{eff}, k_b, k_u) = \int_\Lambda [u(x,t) + v(x,t)]\, dx$$

- The effective diffusion D_{eff} accounts implicitly for a freely diffusing and a weakly bound population
- The model has been used to assess the effect of acetylation on the binding affinity of histone H1 (Carrero et al., 2009, Raghuram et al., 2010).
- A diagram model (not a mathematical one) of the histone H1 dynamics with these two populations is proposed (Raghuram et al., 2009).

A binding model for histone H1

(3 parameters: D_{eff}, k_b, k_u)



Diffusing + Weakly Bound

Strongly Bound

$$\frac{\partial}{\partial t} u(x,t) = D_{eff} \frac{\partial^2}{\partial x^2} u(x,t) - k_b u(x,t) + k_u v(x,t) ,$$

$$\frac{\partial}{\partial t} v(x,t) = k_b u(x,t) - k_u v(x,t)$$

Model drawbacks:

- The proportion of weakly bound population cannot be estimated
- The model does not offer a detailed binding mechanism

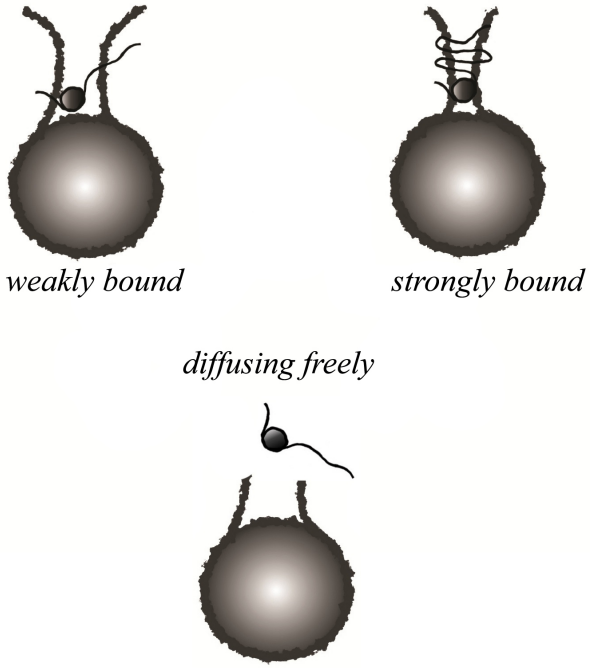
Solution:

- Consider the weakly bound population "explicitly"

An explicit weakly bound population

Three histone H1 populations:

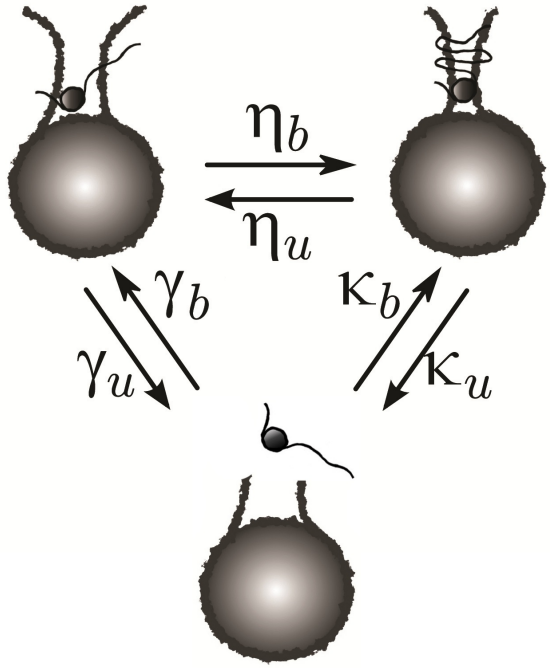
- A freely diffusing population
- A weakly bound population
- A strongly bound population



An explicit weakly bound population

Three histone H1 populations:

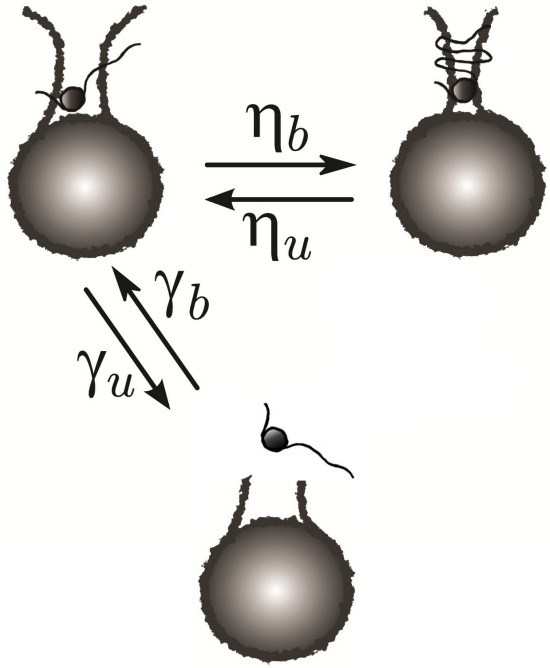
- A freely diffusing population
- A weakly bound population
- A strongly bound population



An explicit weakly bound population

Three histone H1 populations:

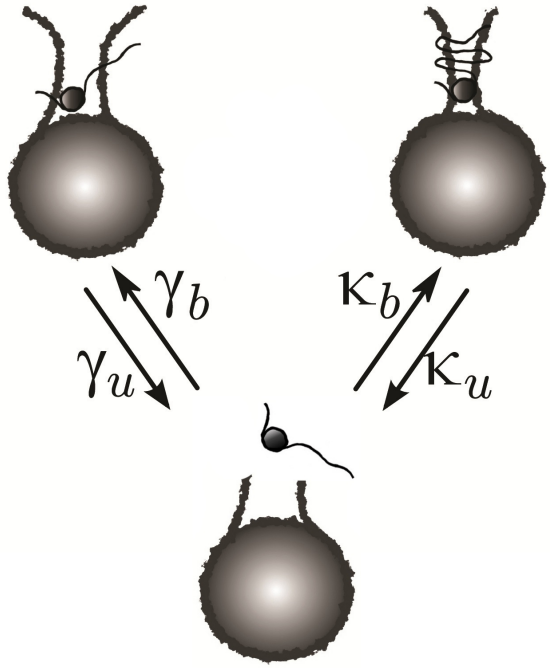
- A freely diffusing population
- A weakly bound population
- A strongly bound population



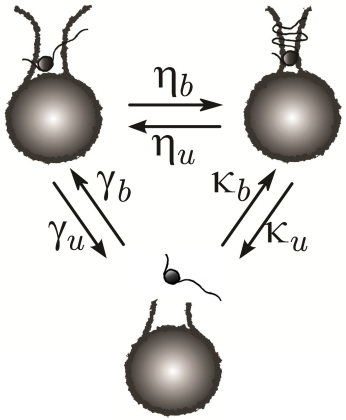
An explicit weakly bound population

Three histone H1 populations:

- A freely diffusing population
- A weakly bound population
- A strongly bound population



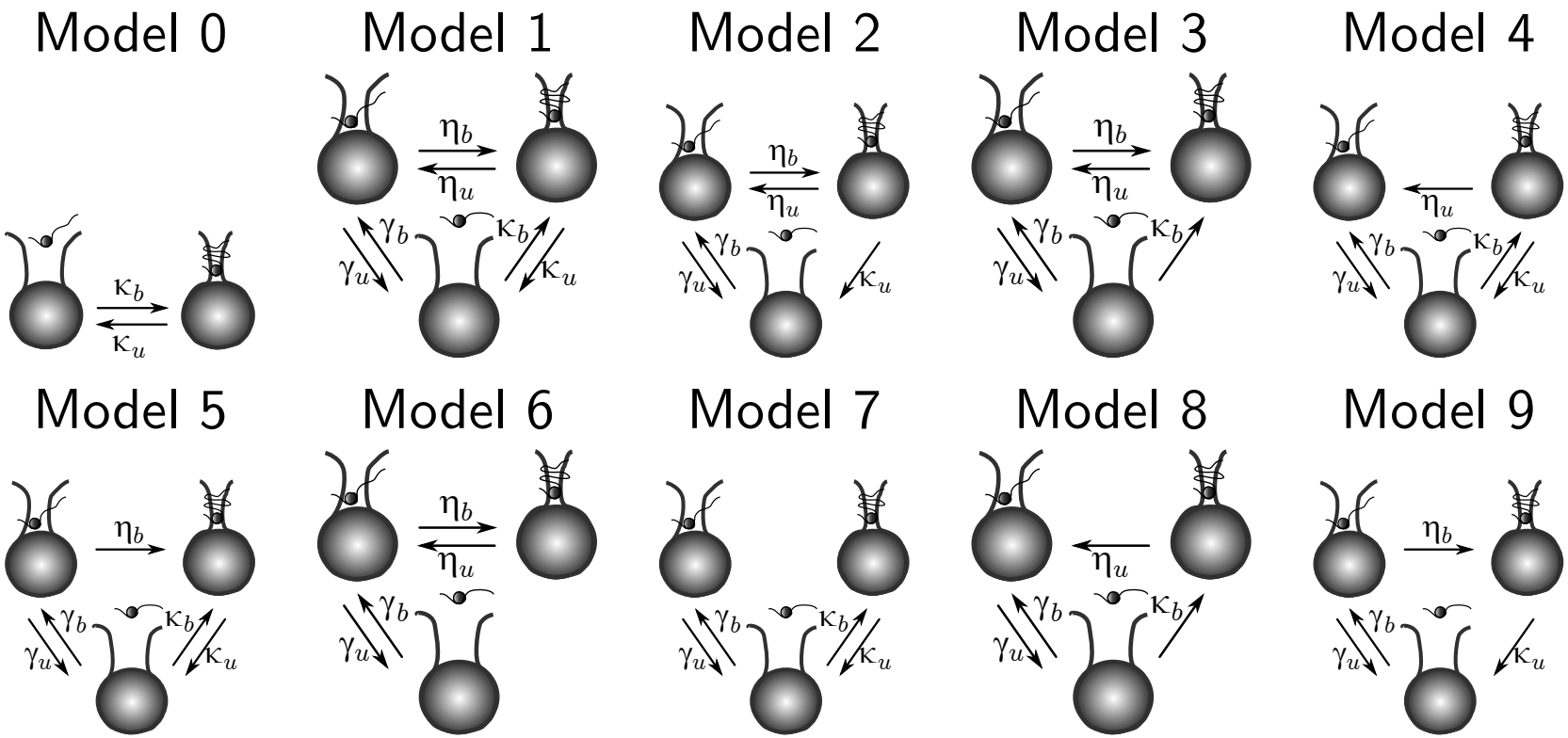
A two binding interaction model for histone H1



$$
\begin{aligned}
u_t &= Du_{xx} - \gamma_b u + \gamma_u w - \kappa_b u + \kappa_u v \\
w_t &= - \eta_b w + \eta_u v + \gamma_b u - \gamma_u w \\
v_t &= + \kappa_b u - \kappa_u v + \eta_b w - \eta_u v
\end{aligned}
$$

- Three populations of histone H1:
  - Unbound: A freely diffusing population u.
  - Weakly bound: A population w weakly bound to the chromatin structure.
  - Strongly bound: A population v strongly bound to the chromatin structure.
- Parameters to be estimated: $D, \kappa_b, \kappa_u, \gamma_b, \gamma_u, \eta_b, \eta_u$.
- Since γ_b and γ_u describe the rapid interaction, they are considered in all the models.

Nested Models



- Some models are nested in others
- Our goal is to select the model that describes experimental data the best

Can we favor one of these mechanisms on the basis of FRAP experiments?

Akaike Information Criterion (AIC)

$$AIC = 2\mathcal{LL}(\hat{p}) - 2n_p = -N \log \frac{\text{RSS}}{N} - 2n_p \ ;$$

- Used when the models are not nested
- *p: set of parameters*
  n_p: number of parameters
  $\mathcal{LL}(\hat{p})$: log of the likelihood $\mathcal{L}(p)$
  $\mathcal{L}(p)$: Likelihood function (probability to find the given data)
  $\hat{p}$: maximum likelihood estimator (parameters that maximize $\mathcal{L}(p)$)
  RSS: sum of square errors (difference of model and observed value)
- The larger the *AIC* the better the model

Likelihood Ratio Test (LRT)

$$H_0: \ M_0 \text{ fits the data well, vs}$$
$$H_1: \ M_1 \text{ fits the data better}$$

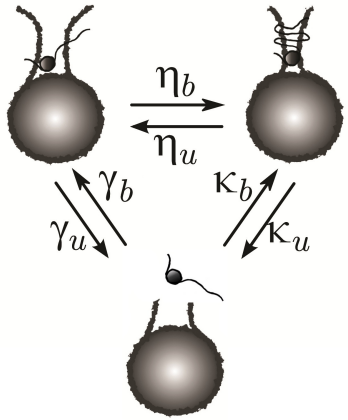
- M_1 is **nested** in model M_0
- The statistic $\lambda = 2(\mathcal{LL}(M_1) - \mathcal{LL}(M_0)) = 2 \log \frac{\mathcal{L}(M_1)}{\mathcal{L}(M_0)}$ is χ^2-distributed
- $\mathcal{LL}(M_0)$ and $\mathcal{LL}(M_1)$: log likelihoods of model M_0 and M_1
- One calculates a χ^2 value given a confidence level and if λ is higher than that value then the null hypothesis is rejected (i.e., M_1 is a better model)

Outline

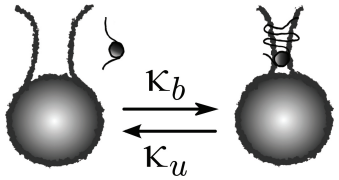
Model Comparison for Nested Models

(general model vs simple model)

M_1:



M_0:

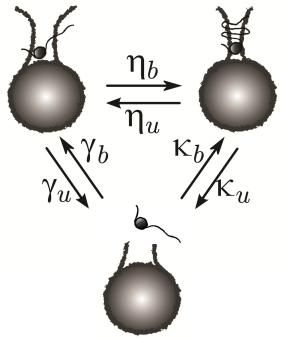


The null hypothesis H_0 is rejected, i.e.,
The general model M_1 is favored using the LRT!

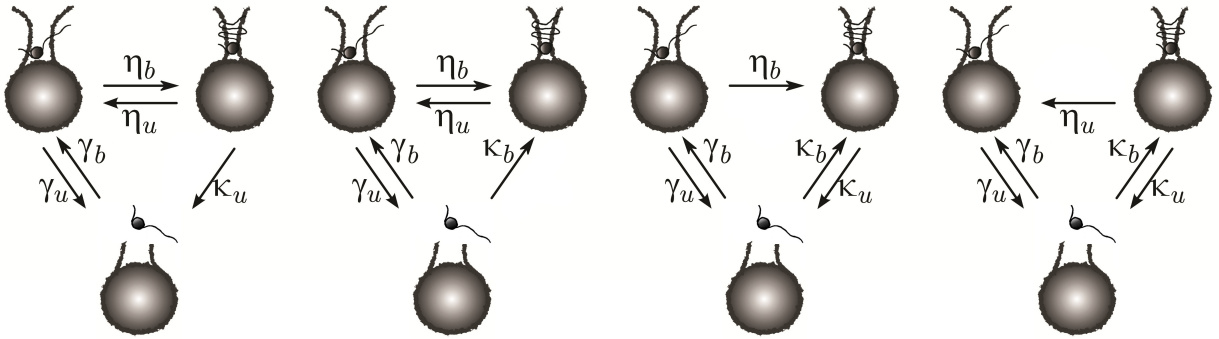
Model Comparison for Nested Models

(general model vs one less interaction models)

M_1:



M_0:

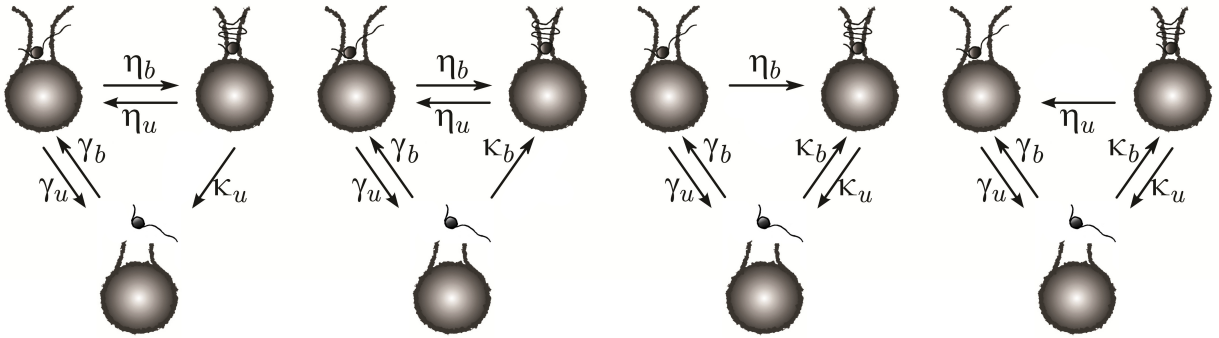


The null hypothesis H_0 cannot be rejected, i.e., all nested models M_0 are favored using the LRT!

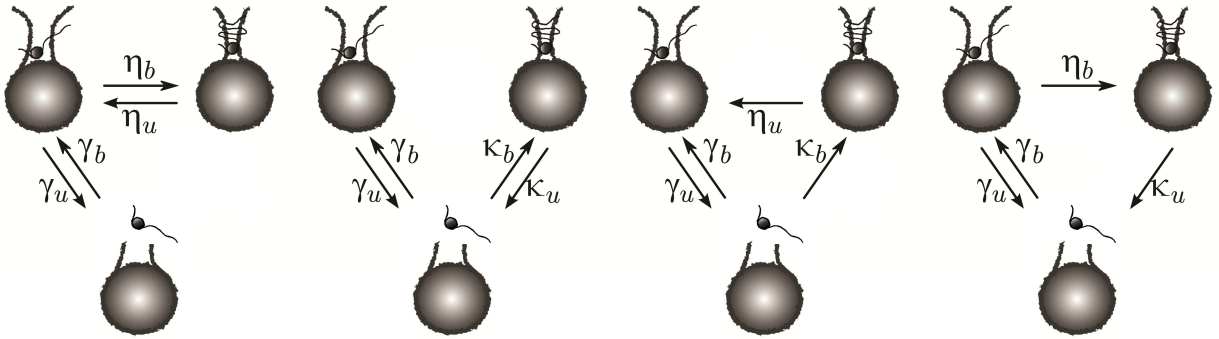
Model Comparison for Nested Models

(one less interaction models vs two less interactions models)

M_1:



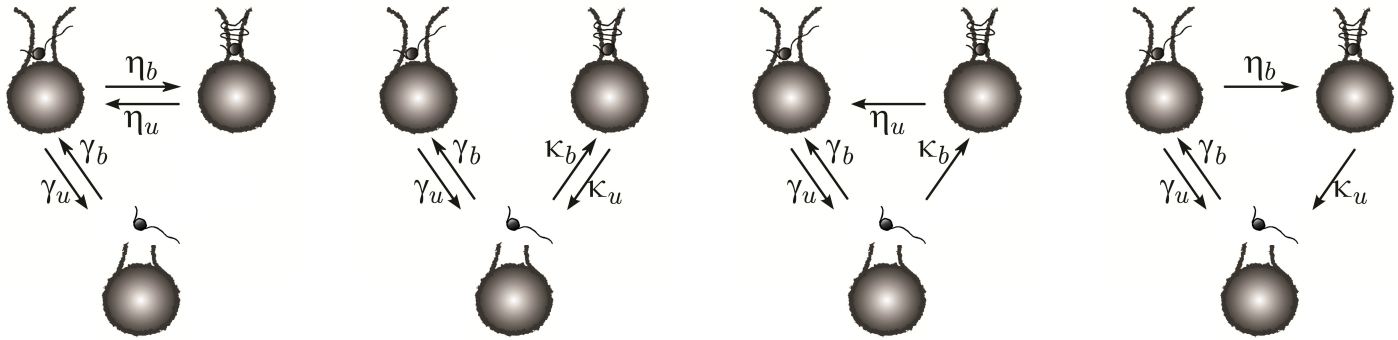
M_0:



The null hypothesis H_0 cannot be rejected, i.e.,
all nested models M_0 with two less interactions are favored using the LRT!

Model Comparison for Non-nested Models

(with two less interactions models)

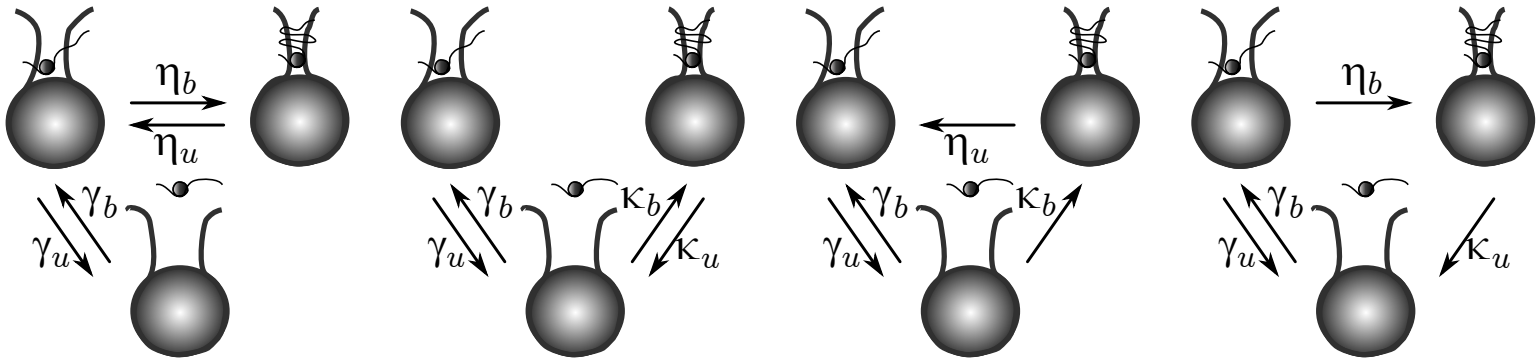


The Akaike Information Criterion (AIC) is not significantly different in any of these models. Thus, none of these models are favored!

Model selection results

Four models showing two types of binding and unbinding processes could describe feasible binding mechanisms of histone H1

- Models 6 to 9
- Four parameters in each model
- Equally feasible according to data



Model inference

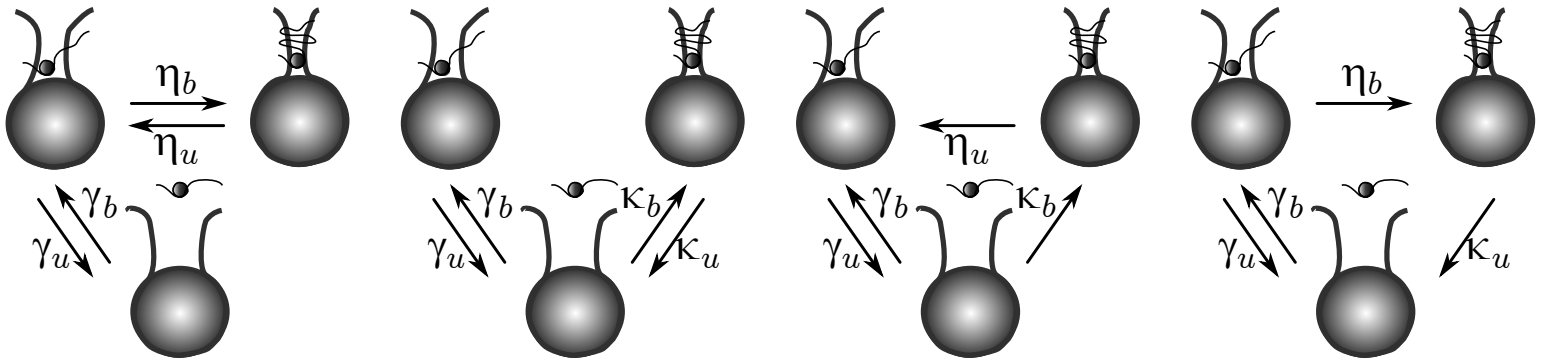
Histone H1.4 data									
Model	P_u	P_w	P_v	$\tau_{u\to v}$	$\tau_{v\to u}$	$\tau_{u\to w}$	$\tau_{w\to u}$	$\tau_{w\to v}$	$\tau_{v\to w}$
$\mathfrak{M}_6$	10%	59%	31%	—	—	11.09s	66.50s	827.56s	433.99s
$\mathfrak{M}_7$	10%	49%	41%	113.42s	474.90s	12.32s	60.89s	—	—
$\mathfrak{M}_8$	10%	54%	36%	129.77s	—	12.13s	60.82s	—	474.54s
$\mathfrak{M}_9$	10%	54%	36%	—	474.61s	11.09s	66.51s	711.82s	—

These models share relevant biological information:

- 10% of the population diffuses across the cell nucleus
- 90% is bound ($\sim$ 55% is weakly bound and 35% strongly bound)
- Same times for rapid interactions

Model inference

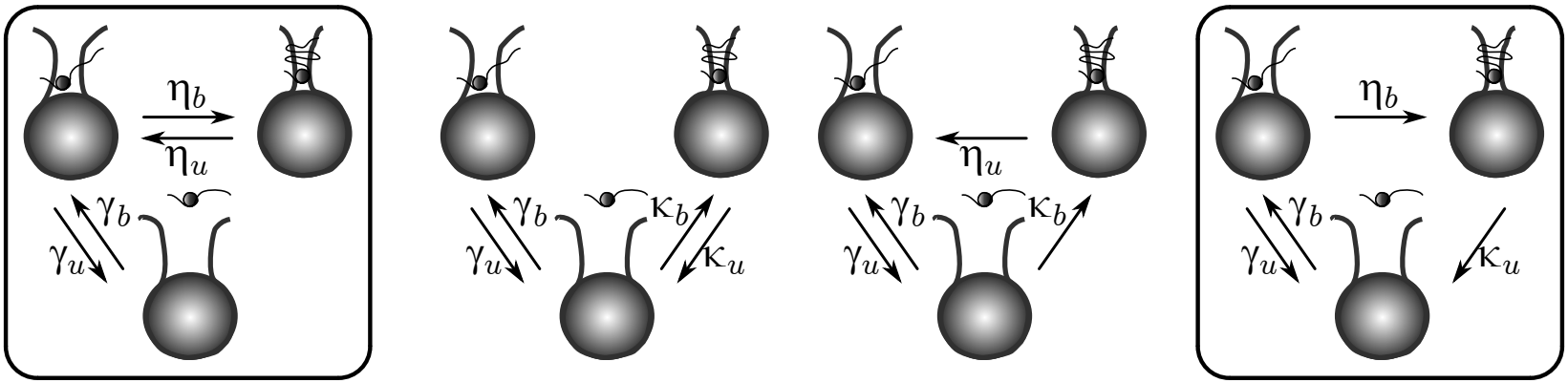
- $\mathfrak{M}_6$ and $\mathfrak{M}_9$: transition to the strongly bound state takes twice as long than unbinding from that state
- $\mathfrak{M}_7$ and $\mathfrak{M}_8$: binding to the strongly bound state (in this case, directly from the free state) is four times faster than unbinding from that state.
- $\mathfrak{M}_6$ – $\mathfrak{M}_9$: they all exhibit the same average residence time in the strongly bound state



Outline

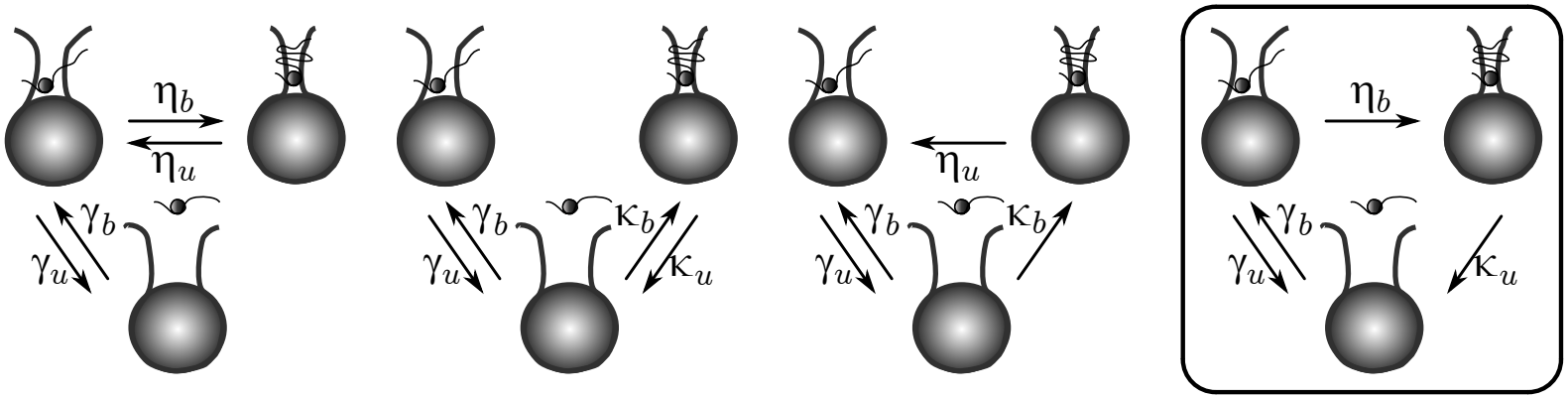
Discussion

- There are two main reasons for ruling out model 6 and 7:
  - A structured binding should require more time than the unbinding
  - Cooperativity suggests transition from weakly to strongly bound (Stasevich et al., 2010)



Discussion

- There are two main reasons for ruling out model 6 and 7:
  - A structured binding should require more time than the unbinding
  - Cooperativity suggests transition from weakly to strongly bound (Stasevich et al., 2010)
- The process for forming a well-defined and stable structure is unlikely to be reversible.



Future Work

- Use the favored model to assess the effect of post-translational modifications on the biding affinity of histone H1 to the chromatin structure.
- Implement the results in a friendly user interface
- Work with data from different histone H1 types
- Analyze the binding mechanism in different regions of the cell nucleus (euchromatin and heterochromatin)

THANK YOU!